ARMY RESEARCH LABORATORY

Empirical Mode Decomposition Based Features for Diagnosis and Prognostics of Systems

**by Hiralal Khatri, Kenneth Ranney,
Kwok Tom, and Romeo del Rosario**

ARL-TR-4301 **April 2008**

Approved for public release; distribution unlimited.

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

Army Research Laboratory
Adelphi, MD 20783-1197

ARL-TR-4301 **April 2008**

Empirical Mode Decomposition Based Features for Diagnosis and Prognostics of Systems

Hiralal Khatri, Kenneth Ranney,
Kwok Tom, and Romeo del Rosario
Sensors and Electron Devices Directorate, ARL

Approved for public release; distribution unlimited.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED (From - To)**
April 2008	Final	

4. TITLE AND SUBTITLE
Empirical Mode Decomposition Based Features for Diagnosis and Prognostics of Systems

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Hiralal Khatri, Kenneth Ranney, Kwok Tom, and Romeo del Rosario

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Research Laboratory
ATTN: AMSRD-ARL-SE-RU
2800 Powder Mill Road
Adelphi, MD 20783-1128

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-TR-4301

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
We present a new procedure to generate additional features for system diagnosis. The procedure is based on empirical mode decomposition of measured signals obtained by monitoring the relevant state of a system. This procedure is different from the existing procedures for defining features, which are generally obtained using the statistics of the measured signal, the matched filter outputs, and the wavelet decomposition of measured signals. Features derived by this new procedure complement the existing features for diagnosis, and therefore they should improve performance of the classifier used to diagnose systems. We illustrate the procedure by generating new features for diagnosis of the AH64A helicopter transmission assembly.

15. SUBJECT TERMS
Empirical mode decomposition, diagnosis, prognostics, features

16. Security Classification of:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Hiralal Khatri
U	U	U	U	30	**19b. TELEPHONE NUMBER** (*Include area code*) (301) 394-0840

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Contents

List of Figures **iv**

Introduction **1**

Empirical Mode Decomposition **1**

Decomposition of Vibration Data **3**

Summary **21**

References **22**

Distribution List **23**

List of Figures

Figure 1. The first 2000 samples of the good data.....3
Figure 2. The first 2000 samples of the faulted data.....4
Figure 3. The power ratios of six data sets relative to the lowest power of the good data set (*4*). 4
Figure 4. The first four empirical mode decompositions from the good system data. 5
Figure 5. The second four empirical mode decompositions from the good system data.....6
Figure 6. The last four empirical mode decompositions from the good system data. 6
Figure 7. The first four empirical mode decomposition from the faulted system data.....7
Figure 8. The second four empirical mode decomposition from the faulted system data. 7
Figure 9. The last four empirical mode decomposition from the faulted system data.....8
Figure 10. Power level for each of the twelve blocks as a function of mode no. Colors denote results of different blocks.....8
Figure 11. Amplitude of the Fourier transform of the first IMF of six data sets. 9
Figure 12. Amplitude of the Fourier transform of the second IMF of six data sets. 10
Figure 13. Amplitude of the Fourier transform of the third IMF of six data sets. 11
Figure 14. Amplitude of the Fourier transform of the fourth IMF of six data sets.....12
Figure 15. Amplitude of the Fourier transform of the fifth IMF of six data sets.....13
Figure 16. Amplitude of the Fourier transform of the sixth IMF of six data sets.....14
Figure 17. Amplitude of the Fourier transform of the seventh IMF of six data sets. 15
Figure 18. Amplitude of the Fourier transform of the eight IMF of six data sets. 16
Figure 19. Amplitude of the Fourier transform of the ninth IMF of six data sets. 17
Figure 20. Amplitude of the Fourier transform of the tenth IMF of six data sets. 18
Figure 21. Ratio of average power of 12 blocks with respect to that of the third good data set as a function of mode number.....19
Figure 22. Ratio of maximum power spectral density of faulted and good data set to the maximum power spectral density of good data set #1. 20
Figure 23. Ratio of maximum power spectral density of faulted and good data set to the power spectral density of good data set #1 at the same frequency where the maximum density occurred for the test data. 21

Introduction

The U.S. Army can enhance availability of its war-fighting equipment and reduce costs by employing modern, condition-based, predictive maintenance. The dynamics of the electromechanical systems used in equipment such as helicopters, tanks, trucks, etc., are quite complex, including nonlinearity, nonstationarity, quasi-periodicity and occasional chaos, especially if there is a mechanical fault. Often, the pertinent parameters of the system cannot be measured conveniently, so the condition of the equipment is estimated on the basis of the vibrations measured at selected locations on the housing of the machinery. Sometimes, certain characteristics of the vibrations can be related to faults in gears or bearings, but in general the relationship between the vibrations and the operational dynamics of the underlying system are unknown or too difficult to characterize. Consequently, empirical algorithms are often developed to relate vibration characteristics and the condition of the equipment by analyzing the measured vibrations under different operating conditions, and sometimes with seeded faults in laboratory experiments.

Since rolling element bearings and gears are widely used in modern rotating machinery, a number of procedures (*1-5*) have been developed to detect faults in such components based on measured vibrations. Algorithms based on statistics of the measured vibration are summarized in (*1*). A number of diagnostic procedures have been developed by using wavelet decomposition (*2-4*). Phase-space based features for fault detection are discussed in (*5*). The vibrations resulting from a fault in rotating machinery will have cyclical behavior and analysis of cyclostationarity is reviewed in (*6*). Fusion of features for predictive diagnosis of electromechanical systems is discussed in (*7*). For successful fusion, we need features that are statistically independent. We present new features based on empirical mode decomposition of vibration data, which may complement existing features and thereby enhance accuracy of diagnosis by fusion of all the features.

Empirical Mode Decomposition

Empirical mode decomposition (EMD), intrinsic mode function (IMF), and the subsequent Hilbert transform for nonlinear and non-stationary time series analysis is described in (*8*) together with numerical procedure illustrated with examples. Since EMD is based on the local characteristic time scale of the data, we want to investigate if it can help generate new features for detection of faults in electromechanical systems. The decomposition of a given time series is based on direct extraction of the energy associated with various intrinsic time scales of the data, and we refer the reader to (*8*) for details of the method. Briefly, local maxima and minima of the

data, *x(t)*, are identified and the maxima are connected by a cubic spline line as the upper envelope. Similarly, a lower envelope is generated from the minima. The two envelopes should bracket all the data between them. Let $E_u(t)$ and $E_l(t)$ denote the values of upper and lower envelopes as a function of discrete time t. The mean of the two envelopes is

$$M(t)=(E_u(t) + E_l(t))/2.$$

The difference between the data and the mean gives the residue:

$$h_1(t)=x(t) – M(t),$$

where the subscript denotes level of iteration. This process is iterated with $h_1(t)$ representing a new time series whose upper and lower envelopes and mean envelope, $M_1(t)$, are computed to obtain the value for next iteration:

$$h_{i+1}(t)=h_i(t)-M_i(t), \text{ where } i=1,2,….$$

Ideally, this process is continued until the residue $h_i(t)$ satisfies the conditions specified for IMF: (1) in the whole data set $h_i(t)$, the number of extrema and the number of zero crossings must either equal or differ at most by 1; and (2) $M_i(t)=0$ for all t. It is not always advisable or efficient to test for these conditions. Instead, the authors (*8*) recommend that the iteration process be terminated when the value of the standard deviation, SD:

$$SD = \sum_{t=0}^{T} \left[\frac{[h_i(t) - h_{i-1}(t)]^2}{h_{i-1}^2(t)} \right],$$

is below a selected value, generally in the range of (0.2-0.3). Note that when $h_i(t) = h_{i-1}(t)$ for all t, no further reduction in residue is possible so $h_i(t)$ is considered to have converged and $SD=0$. The value of SD is used to gage the level of convergence.

This gives us the first IMF, denoted as $c_1(t)=h_i(t)$. To obtain subsequent decompositions, c_i, i=2, 3… treat the residue r_i:

$$r_i=r_{i-1} -c_i(t), \text{ with } r_0=x(t)$$

as the signal to be decomposed and repeat the above procedure to obtain the IMF of the residue. This decomposition process can be stopped based on predetermined conditions. Note that if n IMF's are obtained then

$$x(t) = \sum_{i=1}^{n} c_i(t) + r_n .$$

Decomposition of Vibration Data

We applied the empirical mode decomposition process to the vibration data measured by an accelerometer attached to the underside of the non-rotating main rotor swashplate of a AH64A helicopter (*9*). We had 6 sets of this vibration data, the first two sets were from rotors with faults ("faulted data") and the remaining 4 sets were from rotors with no known faults ("good data"). Each set consisted of 262,144 samples with sampling period of 0.0208 ms, thus, each set represented data collected over 5.4613s. Figure 1 shows the first 2000 samples of one of the good data sets and figure 2 shows similar samples for one of the faulted data sets. Note that the amplitude of the faulted data set is much larger than that of the good data set. Both data sets exhibit some amplitude modulation and there are times when the peaks have negative values and the valleys have positive values. The ratio of the mean squared acceleration for the faulted data sets with respect to the lowest mean squared acceleration obtained from the four good data sets is much higher than the corresponding ratios of the remaining three good data sets; see figure 3. The vibration level of the faulted system is much higher than the good system (ratio near 40) so we may not need any special detector to classify the data set as faulted. However, we will examine how the IMFs for the good and bad data set compare with the view of determining features that may be more sensitive to the condition of the rotor swashplate.

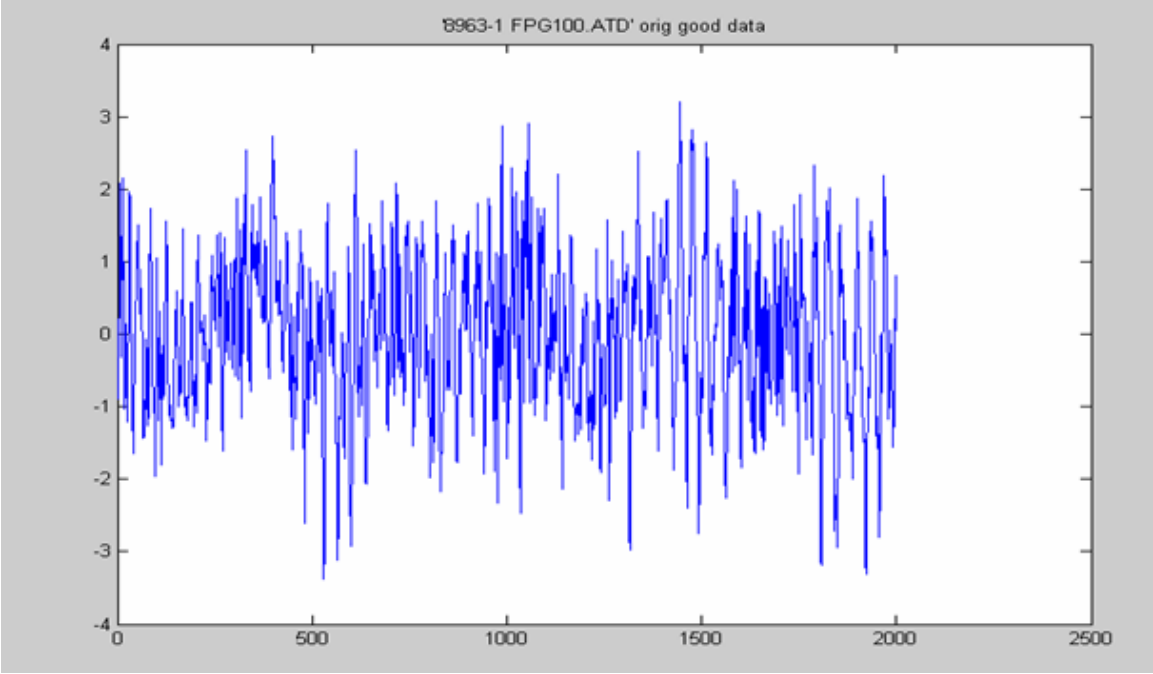

Figure 1. The first 2000 samples of the good data.

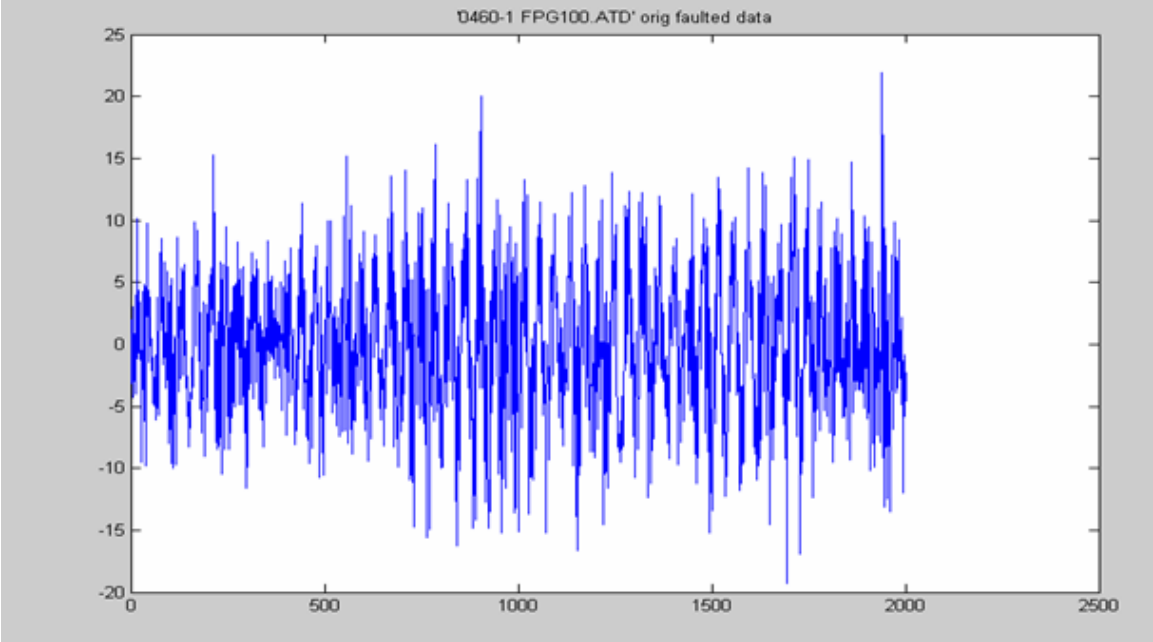

Figure 2. The first 2000 samples of the faulted data**.**

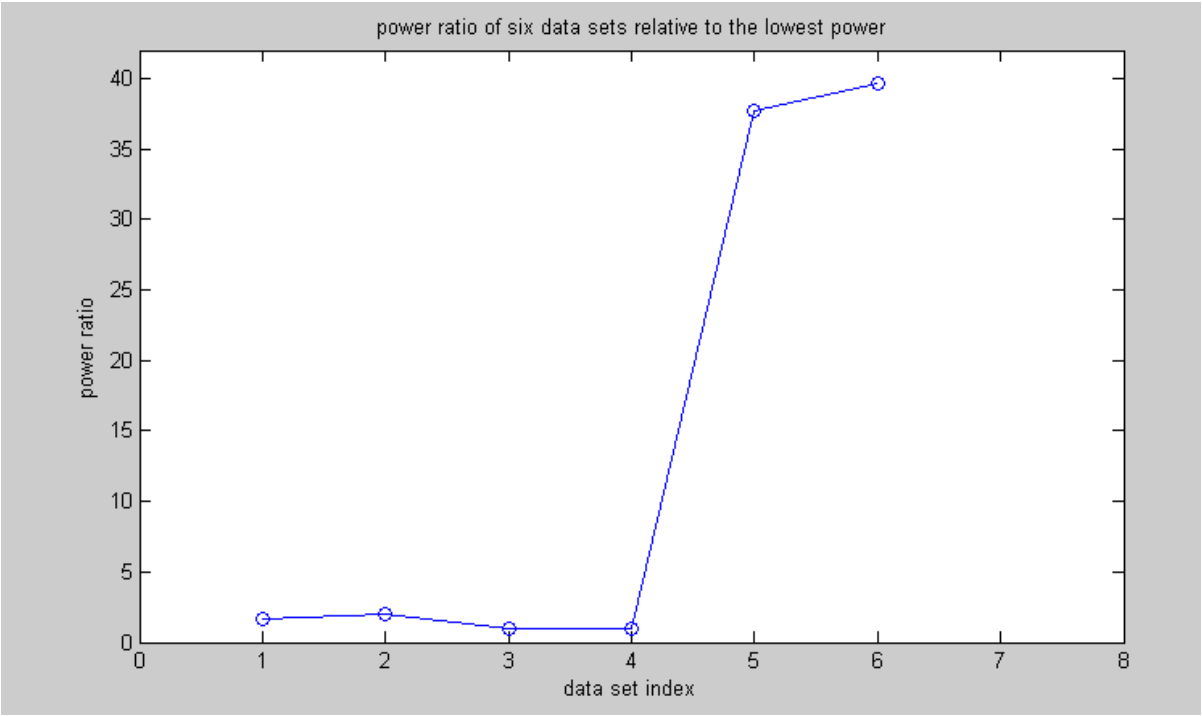

Figure 3. The power ratios of six data sets relative to the lowest power of the good data set (*4*).

We found that the transients of the cubic spline line due to start and end points of the data can significantly alter the values of IMFs. To minimize the effects of the transients, we generated $c_i(t)$s (i=1,…,12) by using all the data points (262,144 samples) for each of the six sets. Each

of the $c_i(t)$s was then separated into 12 consecutive blocks for averaging the statistics of the results. Each block consisted of 20,000 samples and the first block was generated by starting at the 2,001st sample so the first 2,000 samples were discarded to reduce the effects of the spline transients. The 12th block ended at the 242,000th sample so the remaining 22,144 samples were discarded.

The twelve IMFs generated from one of the good data set are shown in figures 4-6. We have plotted only the first 2,000 samples from the first block so that we can show some level of details. Figure 4 shows the first four IMFs starting with c_1 at the top and ending with c_4 at the bottom. Similarly, figure 5 shows c_5 to c_8 and figure 6 shows c_9 to c_{12}. Note that the IMFs are centered about zero and most of the peaks have positive values and most of the valleys have negative values. Figures 7-9 show the corresponding IMFs obtained from one of the faulted data. These IMFs look similar to those of the good data set, except for their amplitudes and frequencies, which we will examine next.

The power in each of the twelve blocks as a function of mode number for one of the good data sets is shown in figure 10. The power level varies slightly from block to block (different colors) for any given mode number but in general, the levels are consistent from one block to another indicating that the blocks have captured the essence of the whole data set.

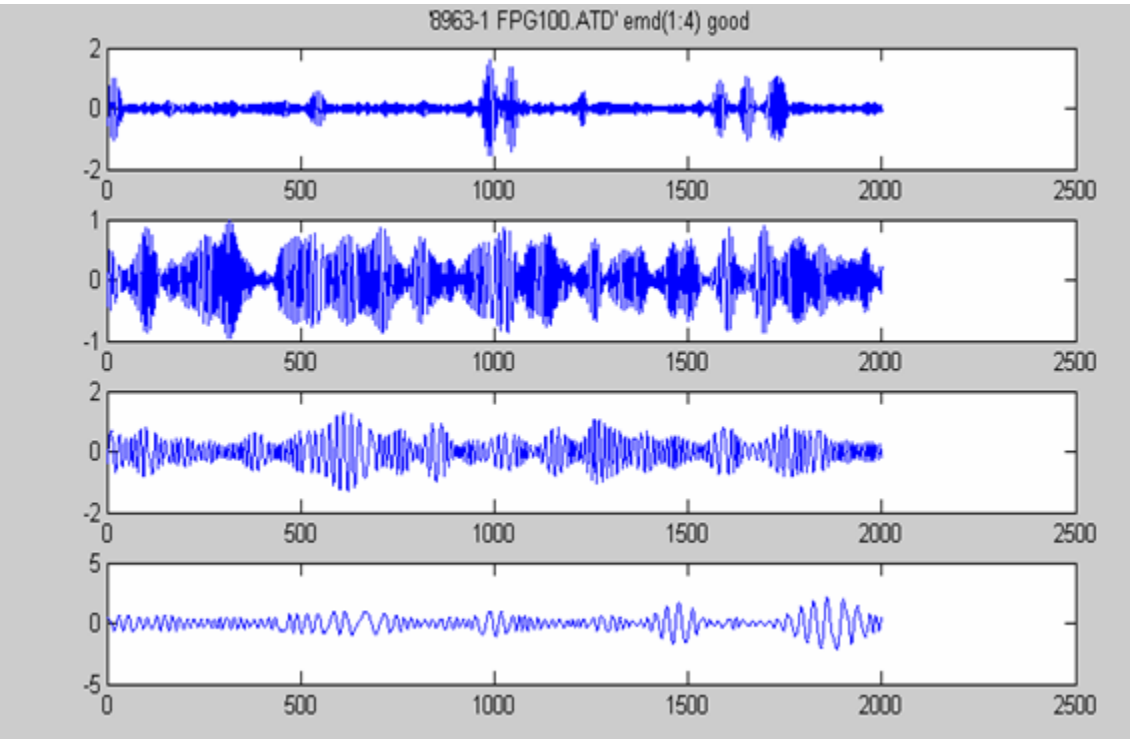

Figure 4. The first four empirical mode decompositions from the good system data.

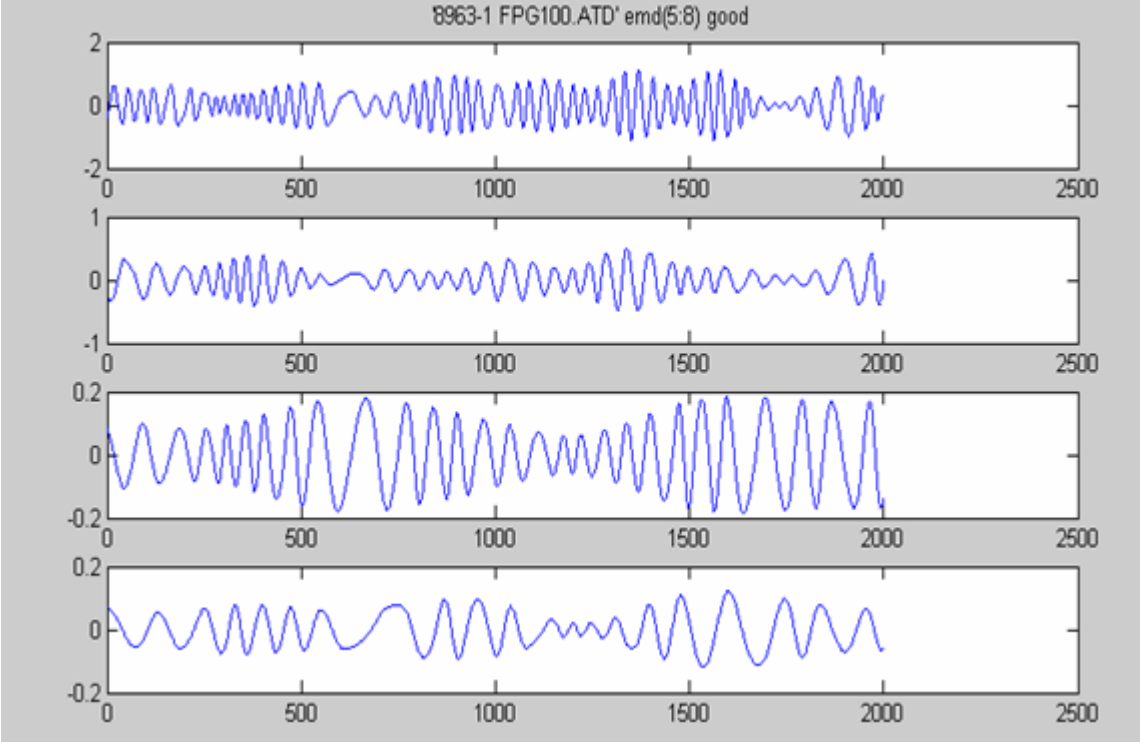

Figure 5. The second four empirical mode decompositions from the good system data**.**

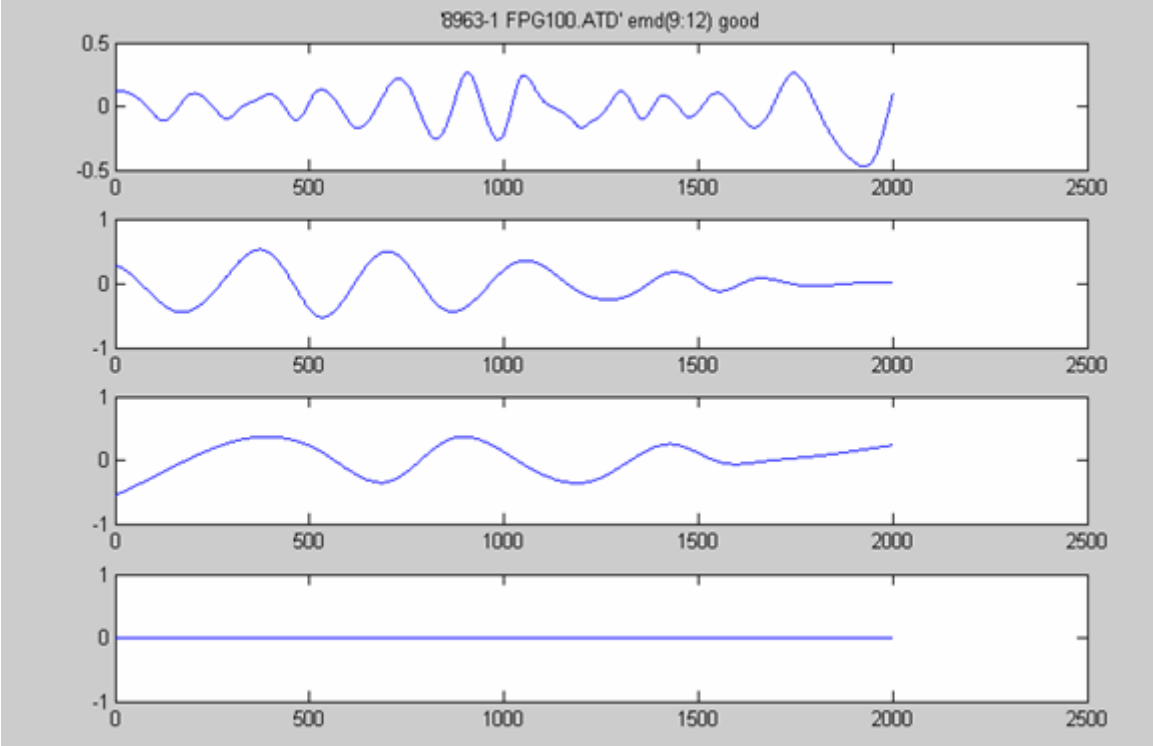

Figure 6. The last four empirical mode decompositions from the good system data**.**

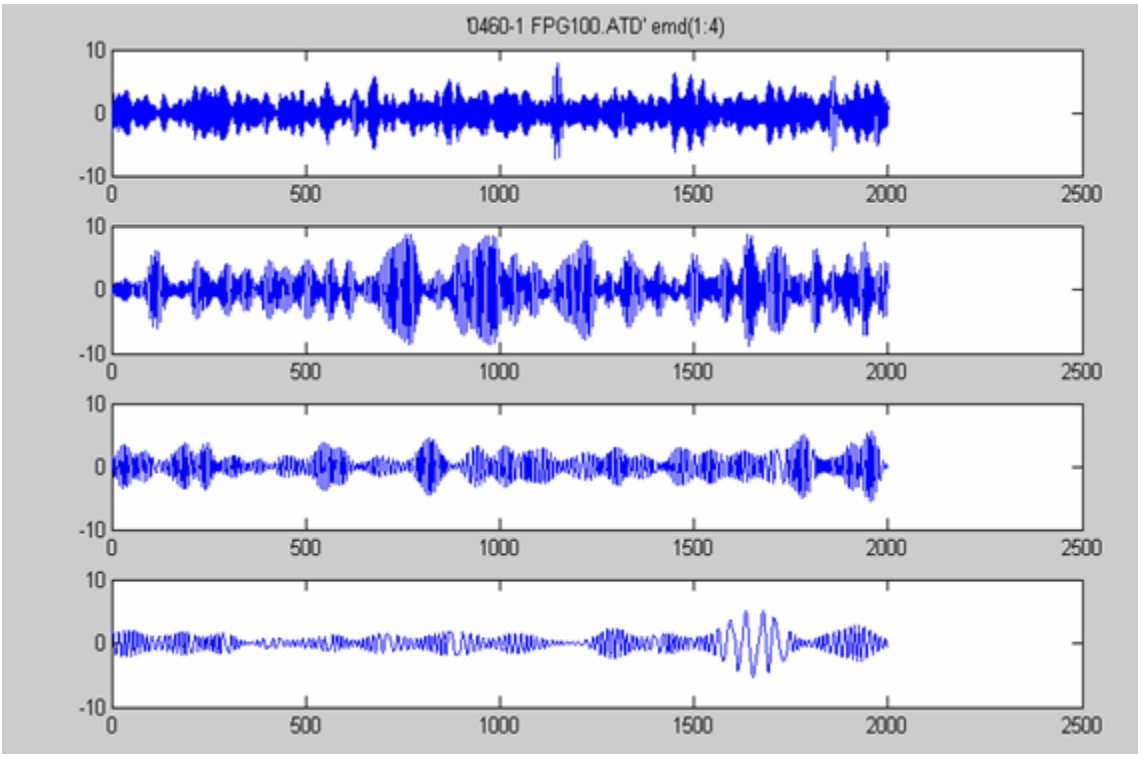

Figure 7. The first four empirical mode decomposition from the faulted system data**.**

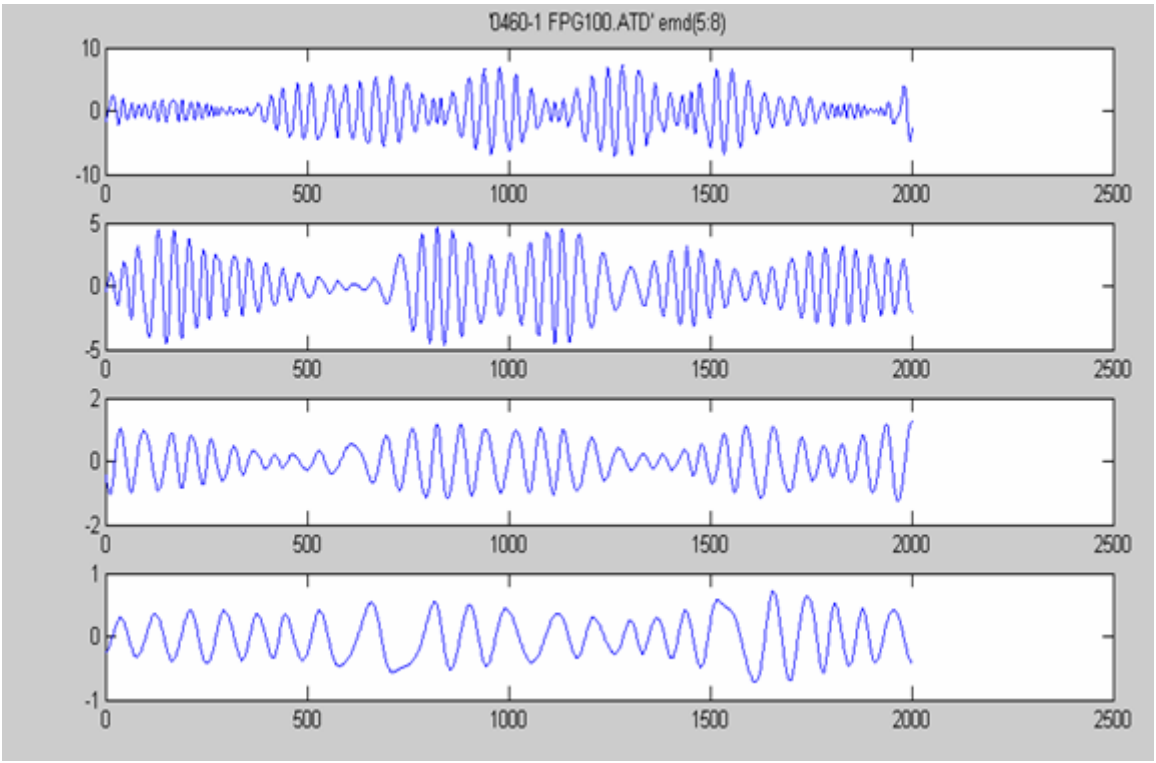

Figure 8. The second four empirical mode decomposition from the faulted system data**.**

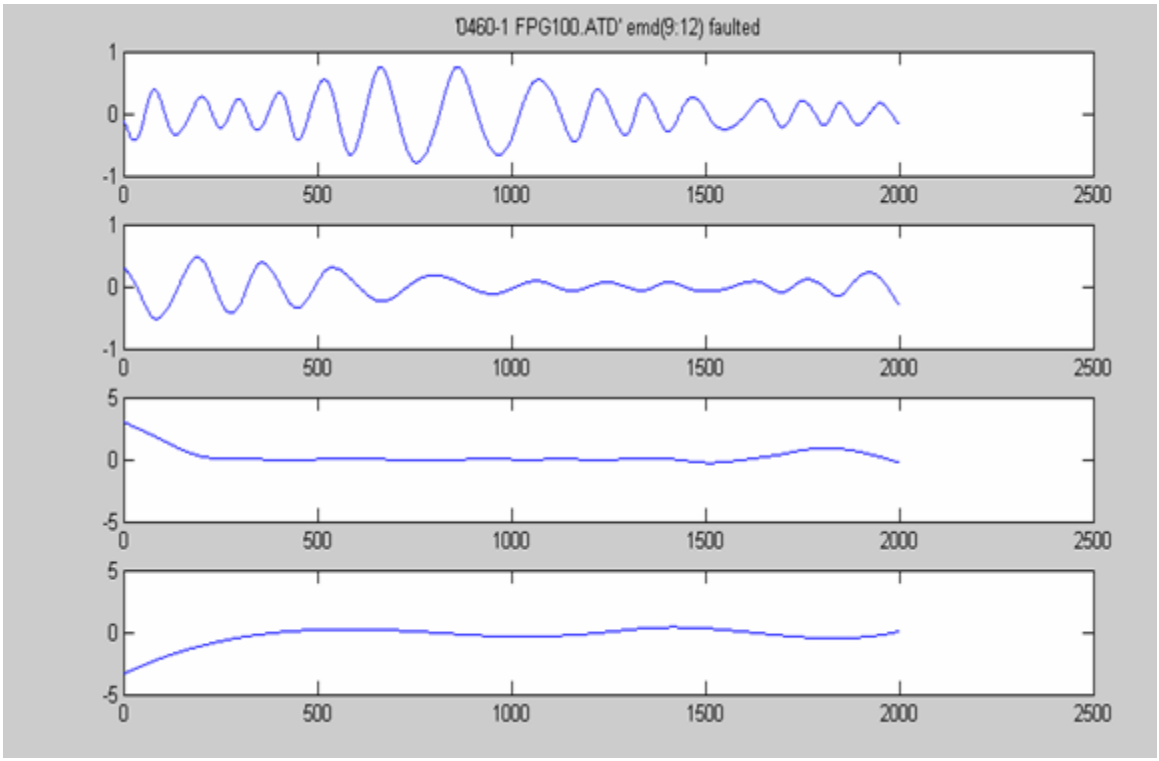

Figure 9. The last four empirical mode decomposition from the faulted system data**.**

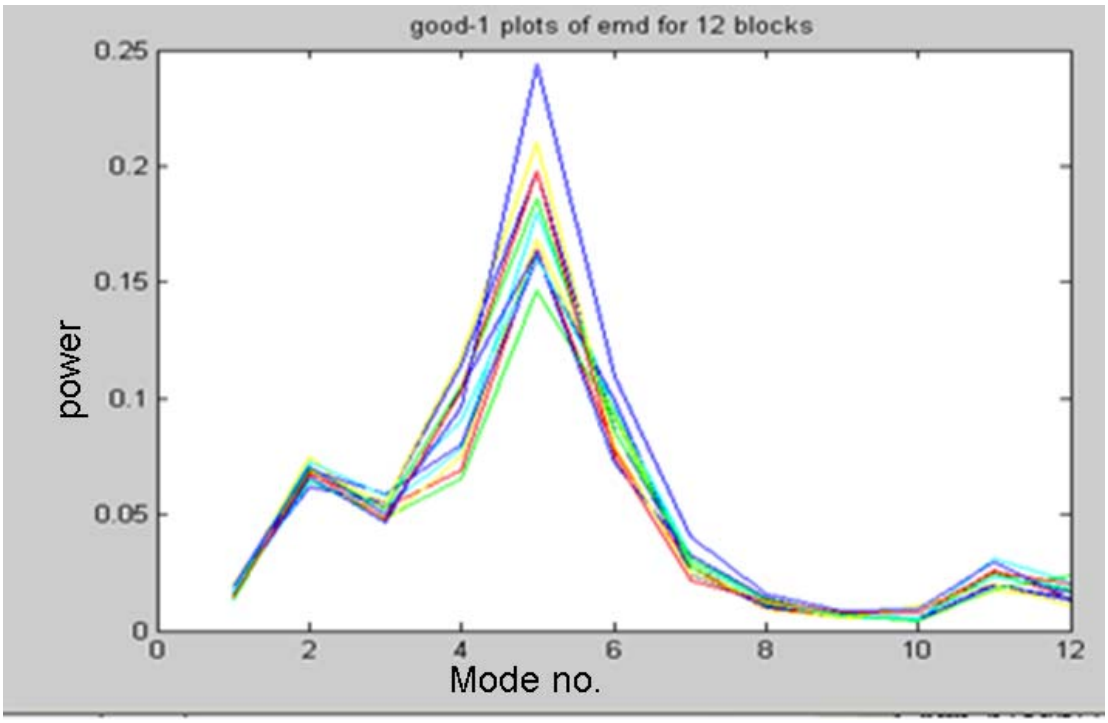

Figure 10. Power level for each of the twelve blocks as a function of mode no. Colors denote results of different blocks.

Next, we examine the Fourier transforms of the IMFs. The Fourier transforms of the IMFs were obtained for each block separately and then the amplitudes were averaged over the twelve blocks. Further, the amplitudes were averaged over 10 frequency samples so that the resultant samples are averaged over 24 Hz. The results are shown in figures 11-20 for the six data sets, two faulted and four good. Only the first 10 modes are shown since the last two modes do not show much variation. Note that the frequencies of the IMFs and the modulation rates decrease as the IMF number increases. There is a difference in frequency where the amplitudes peak and in the frequency bands over which the amplitudes are relatively high for the good and the faulted data sets. We use these differences to generate features for discriminating between the good and faulted data sets.

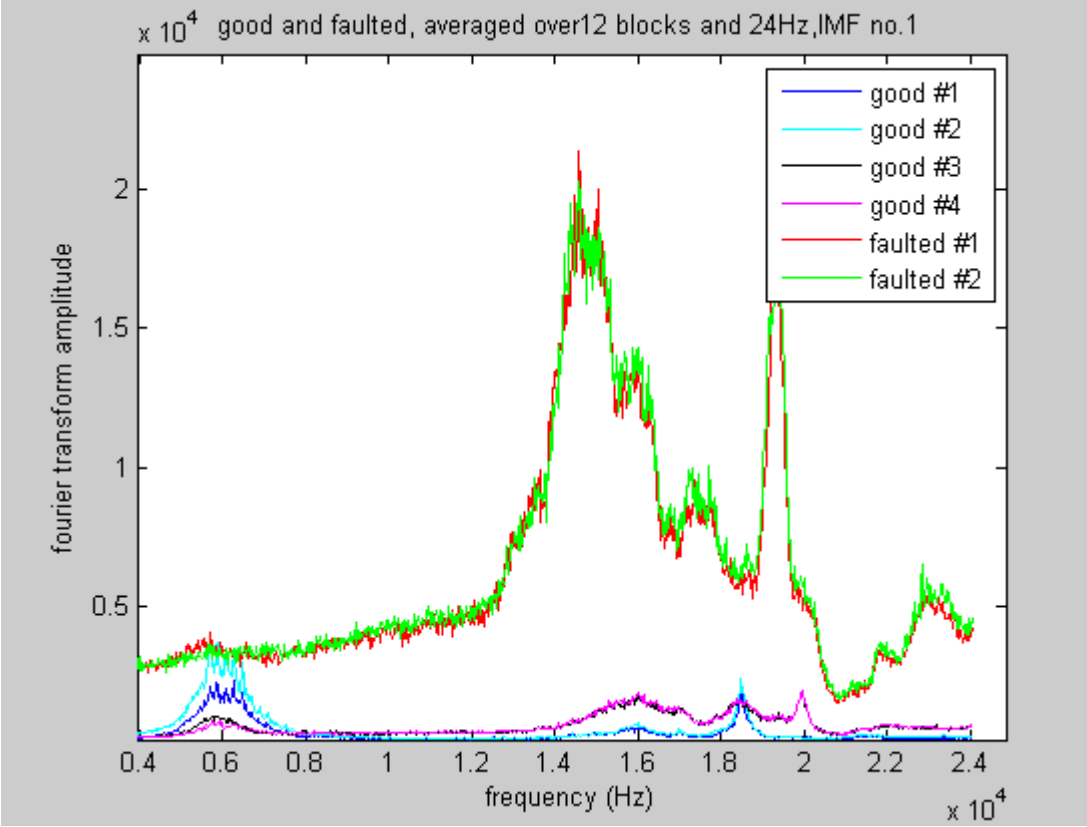

Figure 11. Amplitude of the Fourier transform of the first IMF of six data sets.

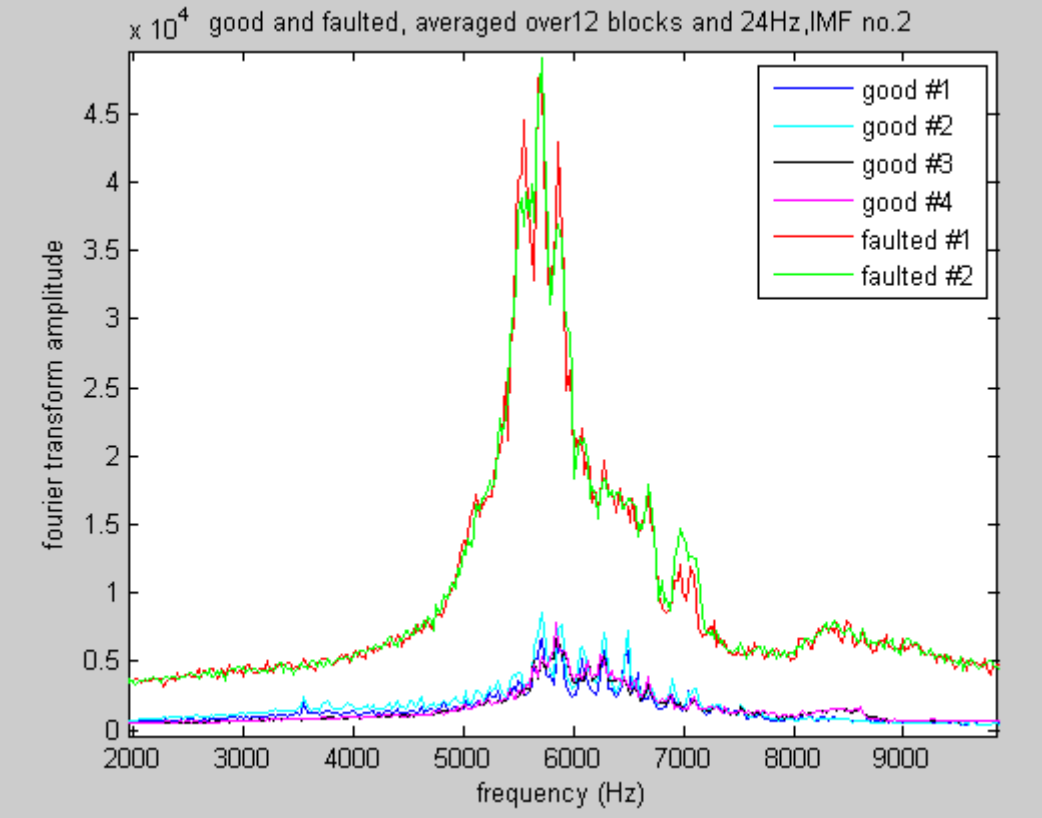

Figure 12. Amplitude of the Fourier transform of the second IMF of six data sets.

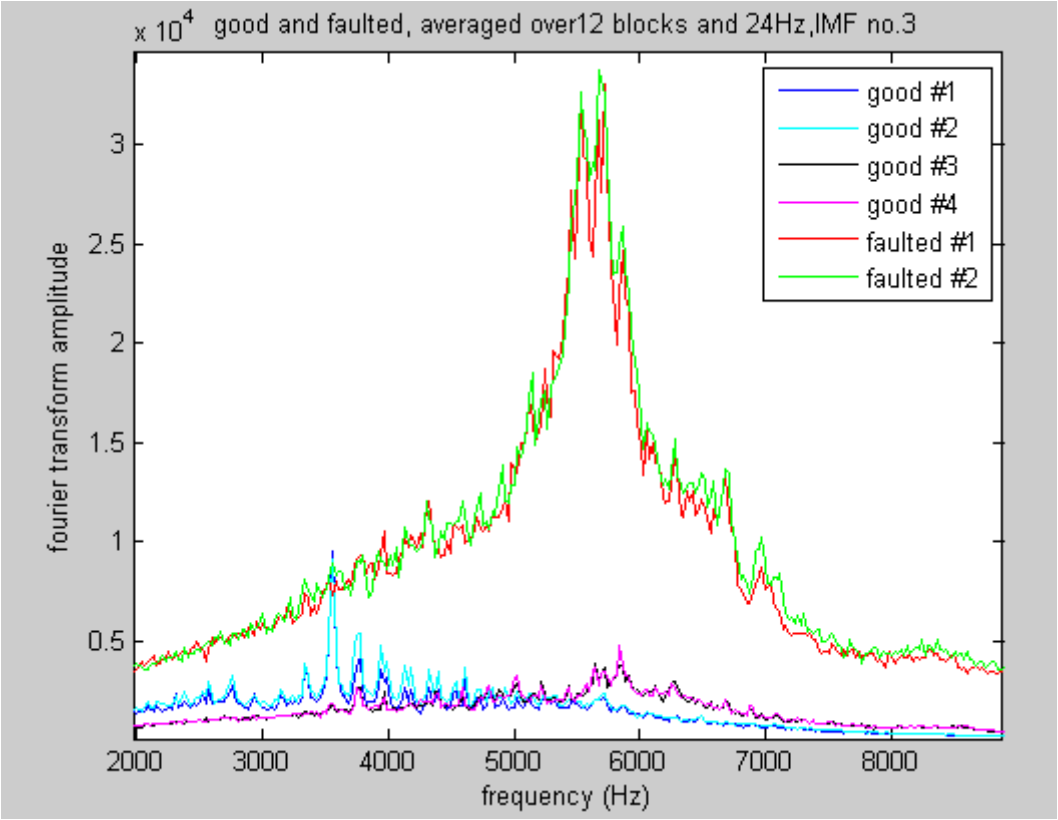

Figure 13. Amplitude of the Fourier transform of the third IMF of six data sets.

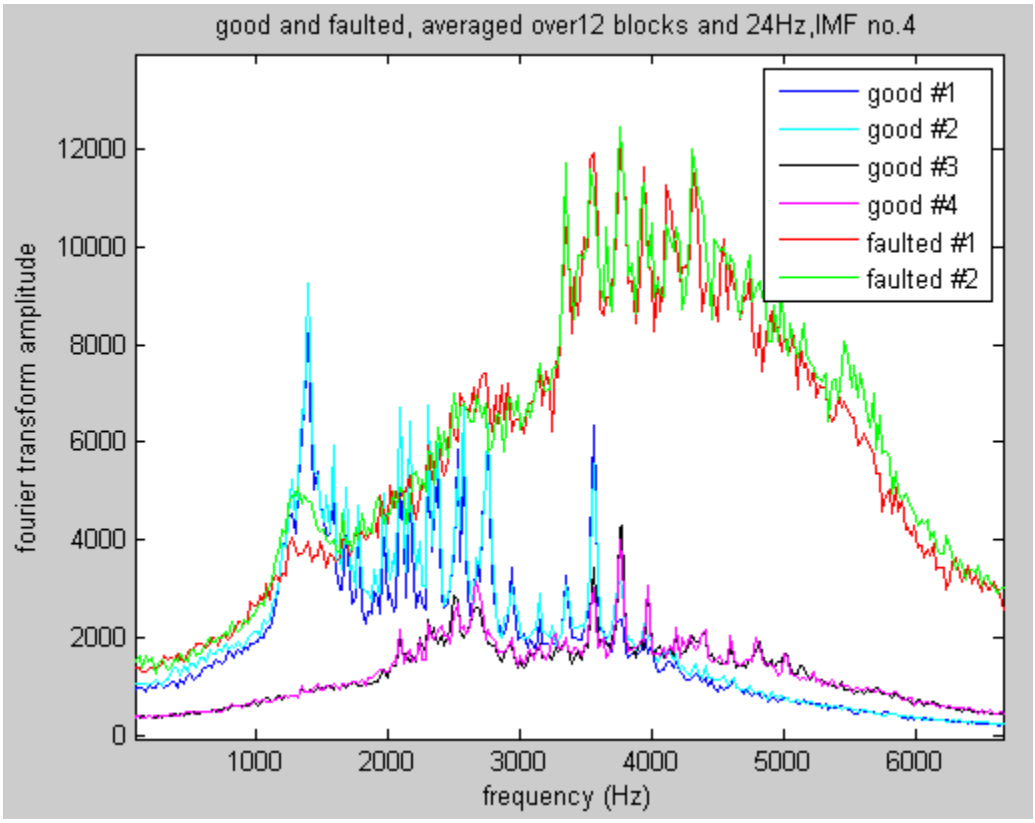

Figure 14. Amplitude of the Fourier transform of the fourth IMF of six data sets.

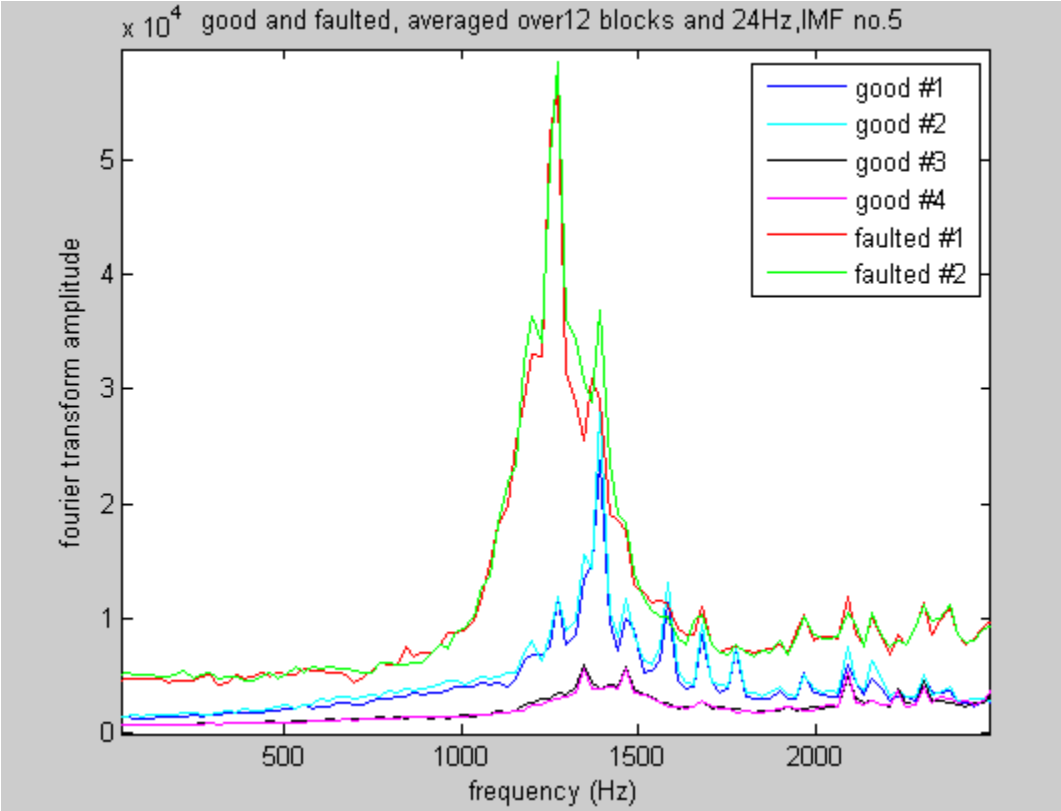

Figure 15. Amplitude of the Fourier transform of the fifth IMF of six data sets.

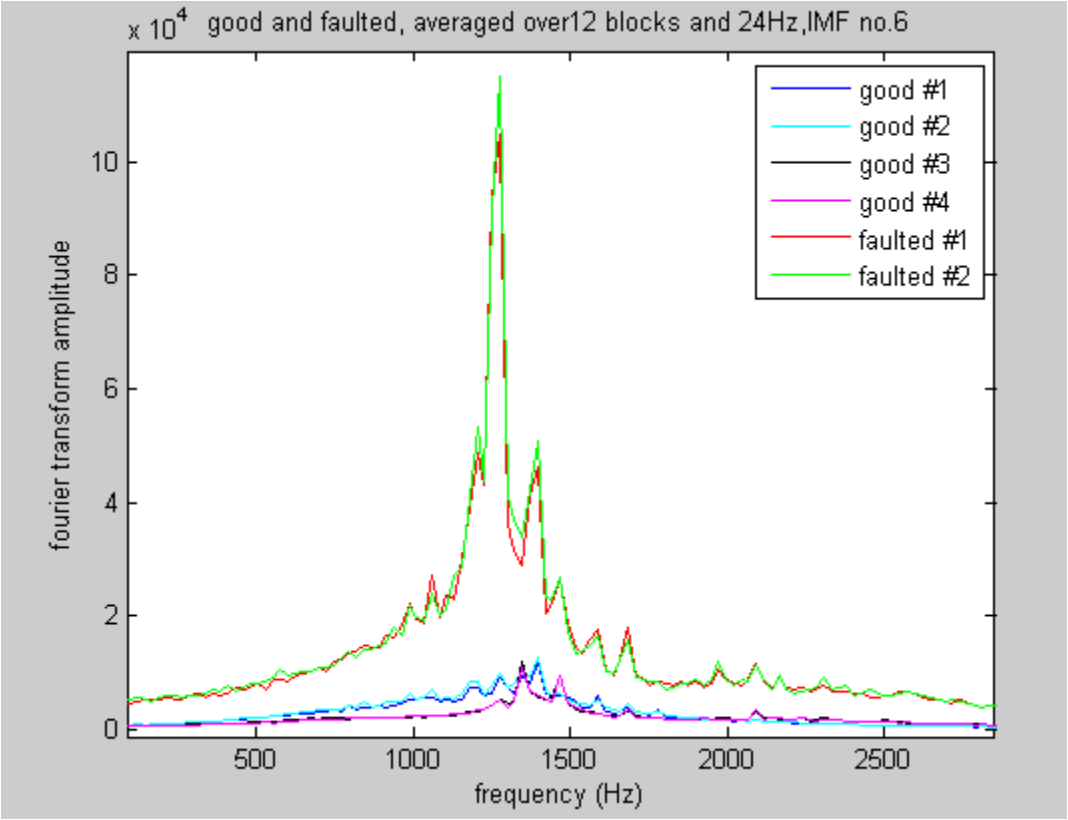

Figure 16. Amplitude of the Fourier transform of the sixth IMF of six data sets.

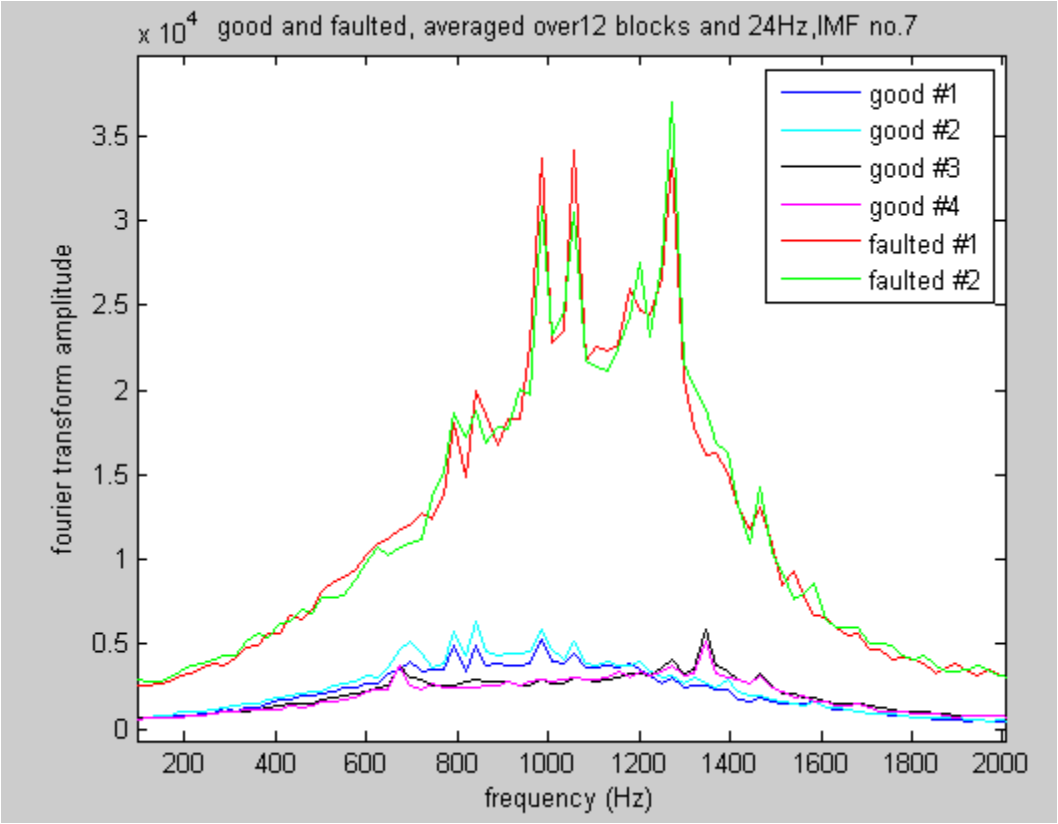

Figure 17. Amplitude of the Fourier transform of the seventh IMF of six data sets.

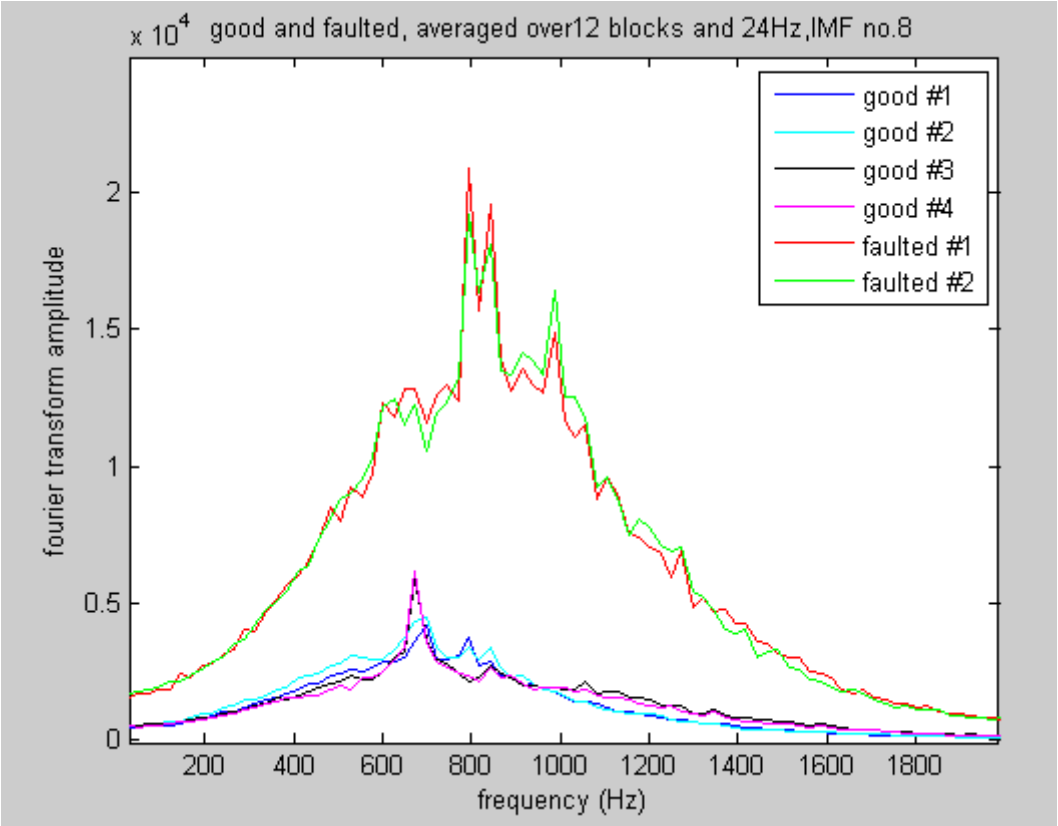

Figure 18. Amplitude of the Fourier transform of the eight IMF of six data sets.

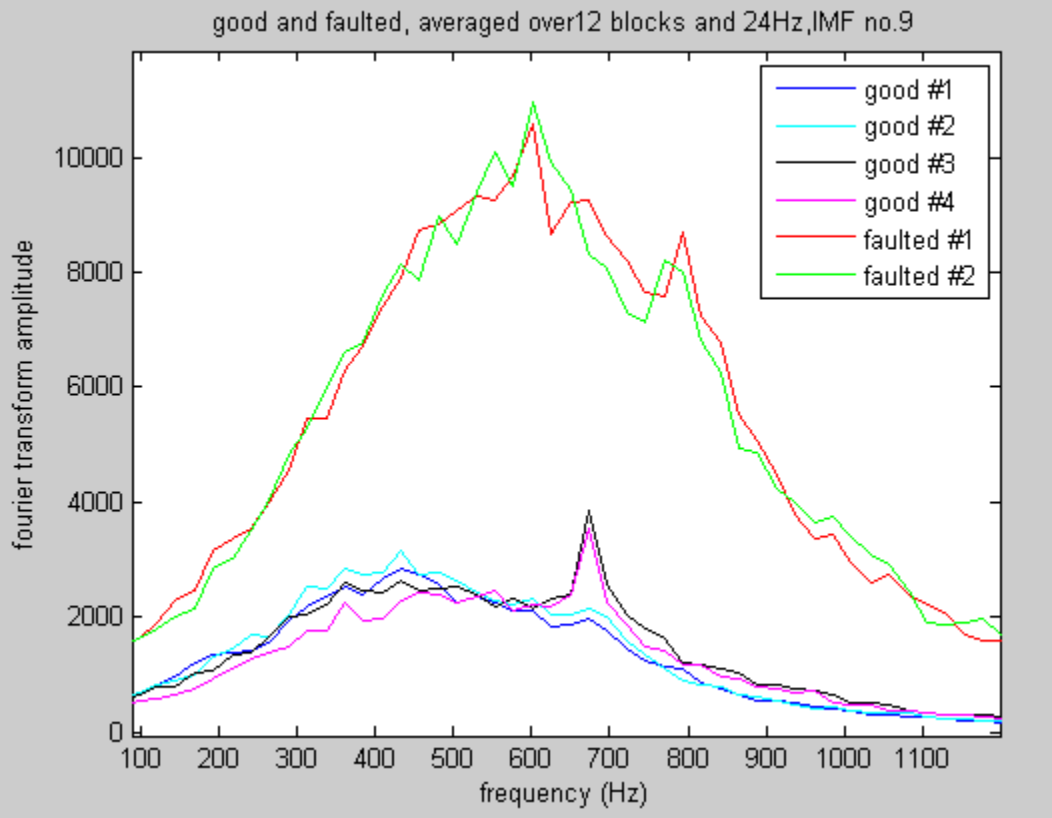

Figure 19. Amplitude of the Fourier transform of the ninth IMF of six data sets.

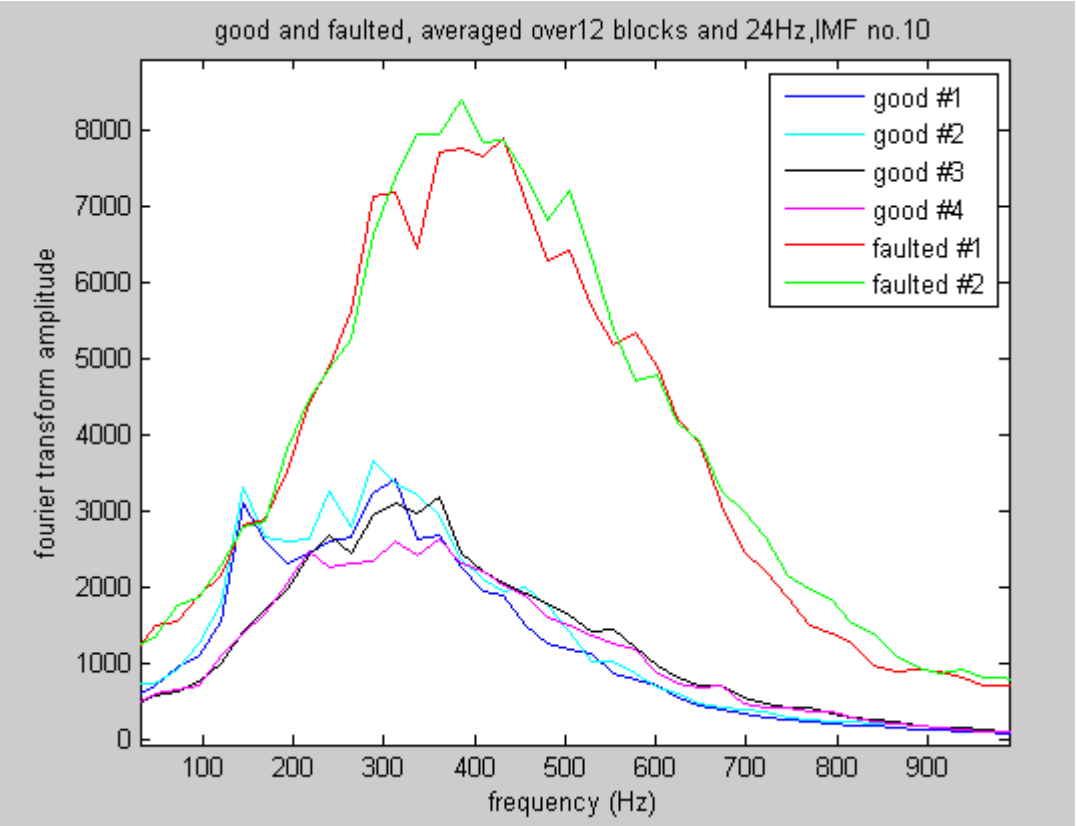

Figure 20. Amplitude of the Fourier transform of the tenth IMF of six data sets.

To gauge variations in powers from one data set to another, we averaged the power over the twelve blocks for each of the modes, and obtained the ratios of the average powers for the two faulted and the three good data sets to the remaining fourth good set for each of the IMFs and the results are shown in figure 21. These ratios for the two faulted data sets are much higher than the ratios for the three good data sets. The ratios for modes 1 thru 8 are greater than 20 for the faulted data sets and less than 5 for the three good data sets. Further, the ratios for mode 1 and 6 of the faulted data sets are greater than 80 and 60, respectively. Thus, the power ratio of these modes are higher than the power ratio of the original time series (near 40) and are therefore better features for classifying faulted and good data. Thus, these ratios for mode 1 and 6 are good candidates for new features.

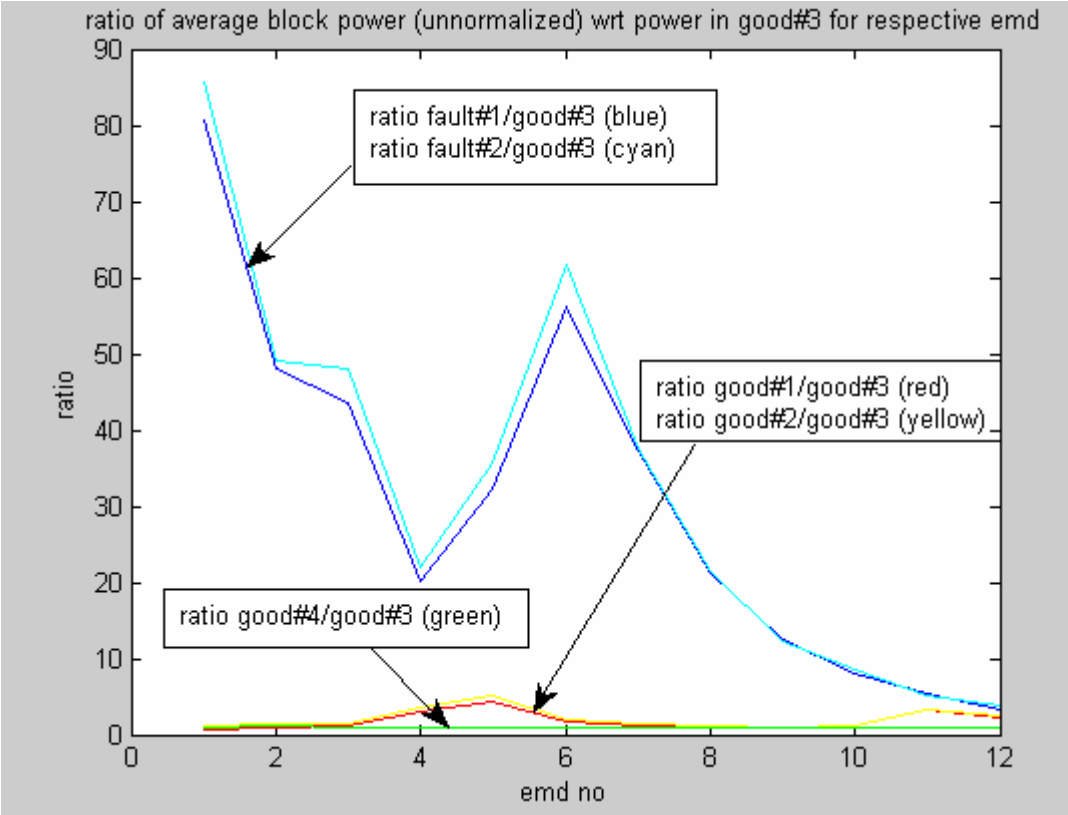

Figure 21. Ratio of average power of 12 blocks with respect to that of the third good data set as a function of mode number.

Next, we took the ratio of the maximum power spectral densities, i.e., the maximum of the squares of the Fourier transforms amplitudes, for all the IMFs of each of the two faulted data sets and three good data sets with respect to the maximum in one of the good data set and the results are given in figure 22. The ratios for IMD no. 1, 2, 6, and 8 of the faulted systems are much higher than for the three good systems and should be potential candidates for the second set of new features to help classify faulted and good data sets. Another candidate for features is shown in figure 23, where we plot the ratio of the maximum power spectral densities to the power density of the one of the good data set (same as the good one used to obtain figure 22) at the same frequency where the maximum density occurred for the test data. This feature takes advantage of the fact that the peaks in power spectrum of the faulted data set are shifted from the corresponding peaks of the good data set. Note that the ratios for modes 1 and 6 are quite high and the reason is obvious from figures 11 and 16. This is our third set of new features.

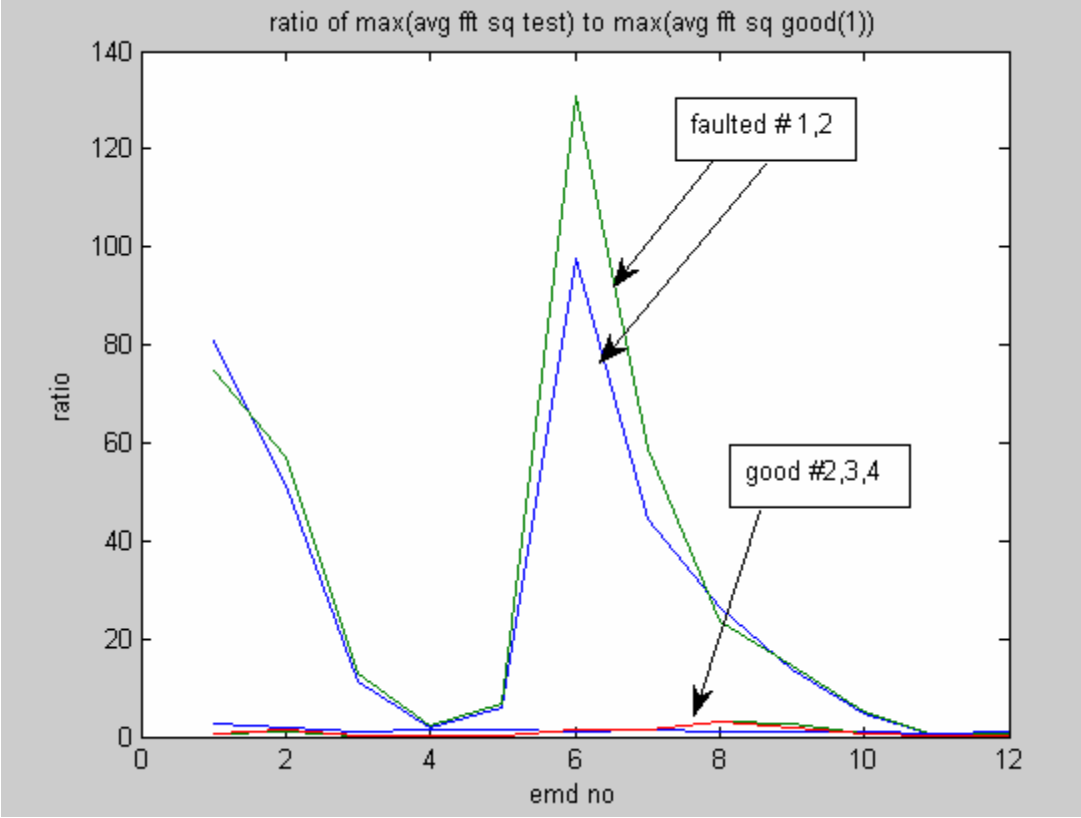

Figure 22. Ratio of maximum power spectral density of faulted and good data set to the maximum power spectral density of good data set #1.

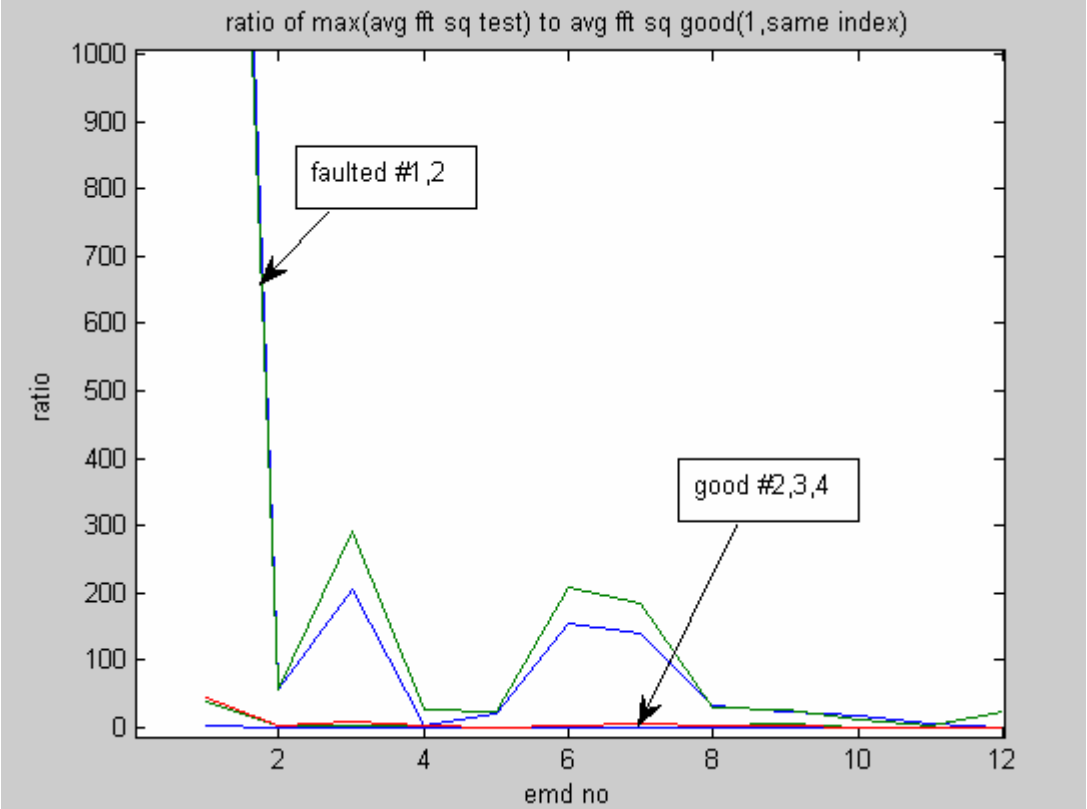

Figure 23. Ratio of maximum power spectral density of faulted and good data set to the power spectral density of good data set #1 at the same frequency where the maximum density occurred for the test data.

We are proposing these sets of three features for assisting in diagnoses of the gear box. Since these features are obtained by using a new methodology, it is likely to complement the existing features and hence, improve the classifier which combines all the features to make a diagnosis.

Summary

Empirical mode decomposition of measured signals of the relevant state of a system can generate additional features for system diagnosis. We developed three sets of new features for diagnosis of transmission box of AH64A helicopter based upon measurements of its vibration data. This procedure is different from the existing procedures for defining features for the transmission box and therefore these new features should complement the existing features for diagnosis and improve performance of the classifier used to diagnose the system.

References

1. Decker, H. J.; Lewicki, D. G. *Spiral Bevel Pinion Crack Detection in a Helicopter Gearbox*; ARL-TR-2958; U.S. Army Research Laboratory: NASA Lewis, June 2003.

2. Tse, P. W.; Peng, Y. H.; Yam, R. Wavelet analysis and envelope detection for rolling element bearing fault diagnosis – their effectiveness and flexibilities. *Journal of Vibration and Acoustics* **July 2001**, ASME.

3. Staszewski, W. J. Structural and mechanical damage detection using wavelet. *The Shock and Vibration Digest* **November 1998**, *30* (6).

4. Wang, C.; Gao, R. Wavelet transform with spectral post-processing for enhanced feature extraction. *IEEE Transactions on Instrumentation and Measurements* **August 2003**, *52* (4).

5. Hively, L. M.; Protopopescu, V. A. Machine failure forewarning via phase-space dissimilarity measures. *Chaos, American Institute of Physics* **June 2004**, *14* (2).

6. Gardner, W. A.; Napolitano, A.; Paura, L. Cyclostationarity: half a century of research. *Signal Processing* **April 2006**, *86* (4).

7. Byington, C.; Garga. Data Fusion for Developing Predictive Diagnostics for Electromechanical Systems, Chapter 23 of CRC Press *Handbook of Multi-sensor Data Fusion*, edited by Hall and Linas, CRC Press, Spring 2001.

8. Huang, N. E.; Shen, Z.; Long, S. R.; Wu, M. C.; Shih, H. H.; Zheng, Q.; Yen, N. C.; Tung, C. C.; Liu, H. H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc., R. Soc. Lond. Ser.* **1998**, *A 454*, 903–995.

9. Keller, J. A.; Branhof, R.; Dunaway, D.; Grabill, P. Examples of condition based maintenance with the vibration management enhancement program. presented at the American Helicopter Society 61st Annual Forum, Grapevine, Texas, June 1–3, 2005.

No. of Copies	Organization
1 (PDF Only)	ADMNSTR DEFNS TECHL INFO CTR ATTN DTIC OCP 8725 JOHN J KINGMAN RD STE 0944 FT BELVOIR VA 22060-6218
4	US ARMY AMRDEC ATTN AMSRD AMR AE A D WADE ATTN AMSRD AMR AE A J KELLER ATTN AMSRD AMR AE KA C RIVERA ATTN AMSRD AMR AE KA E JONES 4484 MARTIN RD RM 247 REDSTONE ARSENAL AL 35898
2	US ARMY ARDEC ATTN AMSRD AAR AEF A M LOSPINUSO ATTN AMSRD AAR AEF A P JONSSON BLDG 91 PICATINNY ARSENAL NJ 07806-5000
1	US ARMY LOGISTICS INNOVATION AGCY (USALIA) ATTN T BRUEN FUTURES GROUP 5870 21ST ST BLDG 212 FT BELVOIR VA 22060-5941
1	US ARMY RSRCH LAB ATTN AMSRD ARL CI OK TP TECHL LIB T LANDFRIED BLDG 4600 ABERDEEN PROVING GROUND MD 21005-5066
3	US ARMY TARDEC ATTN AMSRD TAR N/MS 157 J BECHTEL ATTN AMSRD TAR N/MS 157 J GOTHAMY ATTN AMSRD TAR N/MS 157 T UDVARE 6501 E 11 MILE RD WARREN MI 48397-5000
1	NASA GLENN ATTN AMSRD ARL VT ET D LEWICKI BLDG 23 RM W119A CLEVELAND OH 44135-3191
1	NASA GLENN ATTN AMSRD ARL VT ET H DECKER BLDG 23 RM W121 CLEVELAND OH 44135-3191
1	NASA GLENN ATTN AMSRD ARL VT ET T KRANTZ BLDG 23 RM W115 CLEVELAND OH 44135-3191
1	US GOVERNMENT PRINT OFF DEPOSITORY RECEIVING SECTION ATTN MAIL STOP IDAD J TATE 732 NORTH CAPITOL ST NW WASHINGTON DC 20402
1	CIVIL ENGRG RSRCH FOUNDATION ATTN B MATTES 1015 15TH STRET NW STE 600 WASHINGTON DC 20005
16	US ARMY RSRCH LAB ATTN AMSRD ARL CI OK T TECHL PUB ATTN AMSRD ARL CI OK TL TECHL LIB ATTN AMSRD ARL SE RE K TOM (5 COPIES) ATTN AMSRD ARL SE RE R DEL ROSARIO (5 COPIES) ATTN AMSRD ARL SE RU H KHATRI (2 COPIES) ATTN AMSRD ARL SE RU K RANNEY ATTN IMNE ALC IMS MAIL & RECORDS MGMT ADELPHI MD 20783-1197
Total:	33 (32 HCs, 1 PDF)

INTENTIONALLY LEFT BLANK